Thursday, September 6, 2018

08:45 – 09:00

Registration

09:00 – 10:30

Room: Agora
General Welcome & Invited Talk
Invited Talk Title: *Anonymity in Cryptocurrencies*
Speaker: Dr. Sarah Meiklejohn

10:30 – 11:00

Coffee Break

11:00 – 12:30

Session 1: Privacy Assessment and Trust

Room: C3
Chair: Jorge Cuellar

11:00-11:30 **–Towards an Effective Privacy Impact and Risk Assessment Methodology: Risk Analysis**

By Majed Alshammari and Andrew Simpson (University of Oxford)

Privacy Impact Assessments (PIAs) play a crucial role in providing privacy protection for data subjects and supporting risk management. From an engineering perspective, the core of a PIA is a risk assessment, which typically follows a step-by-step process of risk identification and risk mitigation. In order for a PIA to be holistic and effective, it needs to be complemented by an appropriate privacy risk model that considers legal, organisational, societal and technical aspects. We propose a data-centric approach for identifying and analysing potential privacy risks in a comprehensive manner.

11:30-12:00 - Privacy Risk Assessment: From Art to Science, By Metrics

By Isabel Wagner and Eerke Boiten (De Montfort University)

Privacy risk assessments aim to analyze and quantify the privacy risks associated with new systems. As such, they are critically important in ensuring that adequate privacy protections are built in. However, current methods to quantify privacy risk rely heavily on experienced analysts picking the "correct" risk level on e.g. a five-point scale. In this paper, we argue that a more scientific quantification of privacy risk increases accuracy and reliability and can thus make it easier to build privacy-friendly systems. We discuss how the impact and likelihood of privacy violations can be decomposed and quantified, and stress the importance of meaningful metrics and units of measurement. We suggest a method of quantifying and representing privacy risk that considers a collection of factors as well as a variety of contexts and attacker models. We conclude by identifying some of the major research questions to take this approach further in a variety of application scenarios.

12:00-12:30 - Bootstrapping Online Trust: Timeline Activity Proofs

By Constantin Catalin Dragan and Mark Manulis (University of Surrey)

Establishing initial trust between a new user and an online service, is being generally facilitated by centralized social media platforms, i.e., Facebook, Google, by allowing users to use their social profiles to prove

"trustworthiness" to a new service which has some verification policy with regard to the information that it retrieves from the profiles. Typically, only static information, e.g., name, age, contact details, number of friends, are being used to establish the initial trust. However, such information provides only weak trust guarantees, as (malicious) users can trivially create new profiles and populate them with static data fast to convince the new service. We argue that the way the profiles are used over (longer) periods of time should play a more prominent role in the initial trust establishment. Intuitively, verification policies, in addition to static data, could check whether profiles are being used on a regular basis and have a convincing footprint of activities over various periods of time to be perceived as more trustworthy. In this paper, we introduce Timeline Activity Proofs (TAP) as a new trust factor. TAP allows online users to manage their timeline activities in a privacy-preserving way and use them to bootstrap online trust, e.g., as part of registration to a new service. In our model we do not rely on any centralized social media plat- form. Instead, users are given full control over the activities that they wish to use as part of TAP proofs. A distributed public ledger is used to provide the crucial integrity guarantees, i.e., that activities cannot be tampered with retrospectively. Our TAP construction adopts standard cryptographic techniques to enable authorized access to encrypted activities of a user for the purpose of policy verification and is proven to provide data confidentiality protecting the privacy of user's activities and authenticated policy compliance protecting verifiers from users who cannot show the required footprint of past activities.

12:30 – 14:00

Lunch Break

14:00 – 15:30

Session 2: Private Data and Searches

Room: C3

Chair: Javier Parra-Arnau

14:00-14:30 - Post-processing Methods for High Quality Privacy-preserving Record Linkage

By Martin Franke, Ziad Sehili, Marcel Gladbach, and Erhard Rahm (University of Leipzig)

Privacy-preserving record linkage (PPRL) supports the integration of person-related data from different sources while protecting the privacy of individuals by encoding sensitive information needed for linkage. The use of encoded data makes it challenging to achieve high linkage quality in particular for dirty data containing errors or inconsistencies. Moreover, person-related data is often dense, e.g., due to frequent names or addresses, leading to high similarities for non-matches. Both effects are hard to deal with in common PPRL approaches that rely on a simple threshold-based classification to decide whether a record pair is considered to match. In particular, dirty or dense data likely lead to many multi-links where persons are wrongly linked to more than one other person. Therefore, we propose the use of post-processing methods for resolving multi-links and outline three possible approaches. In our evaluation using large synthetic and real datasets we compare these approaches with each other and show that applying post-processing is highly beneficial and can significantly increase linkage quality in terms of both precision and F-measure.

14:30-15:00 - δ-DOCA: Achieving Privacy in Data Streams

By Bruno C. Leal, Israel C. Vidal, Felipe T. Brito, Juvêncio S. Nobre, and Javam C. Machado (Universidade Federal do Ceará)

Numerous real world applications continuously publish data streams to benefit people in their daily activities. However, these applications may collect and release sensitive information about individuals and lead to serious risks of privacy breach. Differential Privacy (DP) has emerged as a mathematical model to release sensitive information of users while hindering the process of distinguishing individuals' records on databases. Although DP has been widely used for protecting the privacy of individual users' data, it was not designed, in essence, to

provide its guarantees for data streams, since these data are potentially unbounded sequences and continuously generated at rapid rates. Consequently, the noise required to mask the effect of sequences of objects in data streams tend to be higher. In this paper, we design a new technique, named δ-DOCA, to publish data streams under differential privacy. Our approach provides a strategy to determine the sensitivity value of DP and reduces the necessary noise. Our experiments show that the application of δ-DOCA to anonymize data streams not only reduced significantly the necessary noise to apply differential privacy, but also allowed for the output data to preserve the original data distribution.

15:00-15:30 - Data Oblivious Genome Variants Search on Intel SGX

By Avradip Mandal (Fujitsu Laboratories of America), John C. Mitchell (Stanford University), Hart Montgomery (Fujitsu Laboratories of America), and Arnab Roy (Fujitsu Laboratories of America)

We show how to build a practical, private data oblivious genome variants search using Intel SGX. More precisely, we consider the problem posed in Track 2 of the iDash Privacy and Security Workshop 2017 competition, which was to search for variants with high χ2 statistic among certain genetic data over two populations. The winning solution of this iDash competition (developed by Carpov and Tortech) is extremely efficient, but not memory oblivious, which potentially made it vulnerable to a whole host of memory- and cache-based side channel attacks on SGX. In this paper, we adapt a framework in which we can exactly quantify this leakage. We provide a memory oblivious implementation with reasonable information leakage at the cost of some efficiency. Our solution is roughly an order of magnitude slower than the non-memory oblivious implementation, but still practical and much more efficient than naive memory-oblivious solutions–it solves the iDash problem in approximately 5 minutes. In order to do this, we develop novel definitions and models for oblivious dictionary merging, which may be of independent theoretical interest.

15:30 – 16:00

Coffee Break

16:00 – 17:20

Session 3: Internet of Things

Room: C3

Chair: Ken Barker

16:00-16:30 - Developing GDPR Compliant Apps For The Edge

By Tom Lodge, Andy Crabtree, and Anthony Brown (University of Nottingham)

We present an overview of the Databox application development environment or SDK as a means of enabling trusted IoT app development at the network edge. The Databox platform is a dedicated domestic platform that stores IoT, mobile and cloud data and executes local data processing by third party apps to provide end-user control over data flow. Key challenges for building apps in edge environments concern (i) the complexity of IoT devices and user requirements, and (ii) supporting privacy preserving features that meet new data protection regulations. We examine how the Databox SDK can ease the burden of regulatory compliance and be used to sensitize developers to privacy related issues in the very course of building apps.

16:30-17:00 - YaPPL - A Lightweight Privacy Preference Language for Legally Sufficient and Automated Consent Provision in IoT Scenarios

By Max-R. Ulbricht and Frank Pallas (TU Berlin)

In this paper, we present YaPPL — a Privacy Preference Language explicitly designed to fulfill consent-related requirements of the GDPR as well as to address technical givens of IoT scenarios. We analyze what criteria consent must meet in order to be legally sufficient and translate these into a formal representation of consent as

well as into functional requirements that YaPPL must fulfill. Taking into account further nonfunctional requirements particularly relevant in the IoT context, we then derive a specification of YaPPL, which we prototypically implemented in a reusable software library and successfully instantiated in a proof of concept scenario, paving the way for viable technical implementations of legally sufficient consent mechanisms in the IoT.

17:00-17:20 - PrivacyGuard: Enforcing Private Data Usage with Blockchain and Attested Execution
By Ning Zhang (Virginia Polytechnic Institute and State University), Jin Li (Guangzhou University), Wenjing Lou (Virginia Polytechnic Institute and State University), and Y. Thomas Hou (Virginia Polytechnic Institute and State University)
In the upcoming evolution of the Internet of Things (IoT), it is anticipated that billions of devices will be connected to the Internet. Many of these devices are capable of collecting information from individual users and their physical surroundings. They are also capable of taking "smart" actions, which are usually from a backend cloud server in the IoT system. While IoT promises a more connected and smarter world, this pervasive large-scale data collection, storage, sharing, and analysis raise many privacy concerns. In the current IoT ecosystem, IoT service providers have full control of the collected user data. While the original intended use of such data is primarily for smart IoT system and device control, the data is often used for other purposes not explicitly consented to by the users. We propose a novel user privacy protection framework, PrivacyGuard, that aims to empower users with full privacy control of their data. PrivacyGuard framework seamlessly integrates two new technologies, blockchain and trusted execution environment (TEE). By encoding data access policy and usage as smart contracts, PrivacyGuard can allow data owners to control who can have what access to their data, and be able to maintain a trustworthy record of their data usage. Using remote attestation and TEE, PrivacyGuard ensures that data is only used for the intended purposes approved by the data owner. Our approach represents a significant departure from traditional privacy protections which often rely on cryptography and pure software-based secure computation techniques. Addressing the fundamental problem of data usage control, PrivacyGuard will become the cornerstone for free market of private information.

19:00 – 21:00

Social Activity

21:00 – 23:00

Gala Dinner

Friday, September 7, 2018

09:00 – 10:30

Session 4: Privacy and Cryptography

Room: C3
Chair: Guillermo Navarro-Arribas

09:00-09:30 - A Performance and Resource Consumption Assessment of Secret Sharing based Secure Multiparty Computation
By Marcel von Maltitz and Georg Carle (Technical University of Munich)
In recent years, Secure Multiparty Computation (SMC) advanced from a theoretical technique to a practically applicable cryptographic technology. Several frameworks were proposed of which some are still actively developed. We perform a first comprehensive study of performance characteristics of SMC protocols using a

promising implementation based on secret sharing, a common and state-of-the-art foundation. We analyze its scalability with respect to environmental parameters as the number of peers and network properties – namely transmission rate, packet loss, network latency – as parameters and execution time, CPU cycles, memory consumption and amount of transmitted data as variables. Our insights on the resource consumption show that such a solution is practically applicable in intranet environments and – with limitations – in Internet settings.

09:30-10:00 - Privacy-Preserving Trade Chain Detection

By Stefan Wüller (RWTH Aachen University and Stevens Institute of Technology), Malte Breuer (RWTH Aachen University), Ulrike Meyer (RWTH Aachen University), and Susanne Wetzel (Stevens Institute of Technology)

We present a novel multi-party protocol to facilitate the privacy-preserving detection of trade chains in the context of bartering. Our approach is to transform the parties' private quotes into a flow network such that a minimum-cost flow in this network encodes a set of simultaneously executable trade chains for which the number of parties that can trade is maximized. At the core of our novel protocol is a newly developed privacy-preserving implementation of the cycle canceling algorithm that can be used to solve the minimum cost flow problem on encrypted flow networks.

10:00-10:30 - FHE-compatible Batch Normalization for Privacy Preserving Deep Learning

By Alberto Ibarrondo and Melek Önen (EURECOM)

Deep Learning has recently become very popular thanks to major advances in cloud computing technology. However, pushing Deep Learning computations to the cloud poses a risk to the privacy of the data involved. Recent solutions propose to encrypt data with Fully Homomorphic Encryption (FHE) enabling the execution of operations over encrypted data. Given the serious performance constraints of this technology, recent privacy preserving deep learning solutions aim at first customizing the underlying neural network operations and further apply encryption. While the main neural network layer investigated so far is the activation layer, in this paper we study the Batch Normalization (BN) layer: a modern layer that, by addressing internal covariance shift, has been proved very effective in increasing the accuracy of Deep Neural Networks. In order to be compatible with the use of FHE, we propose to reformulate batch normalization, which results in a moderate decrease on the number of operations. Furthermore, we devise a re-parametrization method that allows the absorption of batch normalization by previous layers. We show that whenever these two methods are integrated during the inference phase and executed over FHE-encrypted data, there is a significant performance gain with no loss on accuracy. We also note that this gain is valid both in the encrypted and unencrypted domains.

10:30 – 11:00

Coffee Break

11:00 – 12:20

Session 5: Future Internet

Room: C3

Chair: Isabel Wagner

11:00-11:20 - A general algorithm for *k*-anonymity on dynamic databases

By Julián Salas (Universitat Oberta de Catalunya) and Vicenç Torra (University of Skövde)

In this work we present an algorithm for k-anonymization of datasets that are changing over time. It is intended for preventing identity disclosure in dynamic datasets via microaggregation. It supports adding, deleting and updating records in a database, while keeping k-anonymity on each release. We carry out experiments on database anonymization. We expected that the additional constraints for k-anonymization of dynamic databases would entail a larger information loss, however it stays close to MDAV's information loss for static databases.

Finally, we carry out a proof of concept experiment with directed degree sequence anonymization, in which the removal or addition of records, implies the modification of other records.

11:20-11:40 - On Security of Anonymous Invitation-Based System

By Naoto Yanai and Jason Paul Cruz (Osaka University)

In an anonymous invitation-based system, a user can join a group by receiving invitations sent by current members, i.e., inviters, to a server anonymously. This kind of system is suitable for social networks, and a formal framework with the anonymity of inviters and the unforgeability of an invitation letter was proposed in DPM 2017. The main concept of this previous system is elegant, but the formal security definitions are insufficient and weak in a realistic application scenario. In this paper, we revise formal security definitions as attacks representing a realistic scenario. In addition, we define a new aspect of the security wherein an adversary maliciously generates an invitation letter, i.e., invitation opacity, and the security for guaranteeing that an invitee with a valid invitation letter can always join the system, i.e., invitation extractability. A secure and useful construction can be expected by satisfying the security definitions described above.

11:40-12:00 - Probabilistic metric spaces for privacy by design machine learning algorithms: modeling database changes

By Vicenç Torra (University of Skövde) and Guillermo Navarro-Arribas (Universitat Autònoma de Barcelona)

Machine learning, data mining and statistics are used to analyze the data and to build models from them. Data privacy for big data needs to find a compromise between data analysis and disclosure risk. Privacy by design machine learning algorithms need to take into account the space of models and the relationship between the data that generates the models and the models themselves. In this paper we propose the use of probabilistic metric spaces for comparing these models.

12:00-12:20 - Lifelogging Protection Scheme for Internet-based Personal Assistants

By David Pàmies-Estrems (Universitat Rovira i Virgili), Nesrine Kaaniche (Télécom SudParis), Maryline Laurent (Télécom SudParis), Jordi Castellà-Roca, (Universitat Rovira i Virgili), and Joaquin Garcia-Alfaro (Télécom SudParis)

Internet-based personal assistants are promising devices combining voice control and search technologies to pull out relevant information to domestic users. They are expected to assist in a smart way to household activities, such as scheduling meetings, finding locations, reporting of cultural events, sending of messages and a lot more. The information collected by these devices, including personalized lifelogs about their corresponding users, is likely to be stored by well-established Internet players related to web search engines and social media. This can lead to serious privacy risks. The issue of protecting the identity of domestic users and their sensitive data must be tackled at design time, to promptly mitigate privacy threats. Towards this end, this paper proposes a protection scheme that jointly handles the aforementioned issues by combining log anonymization and sanitizable signatures.

12:20 – 14:00

Farewell & Lunch